RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/544,180A
Source: TFW16
Date Processed by STIC: 11/20/2006

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 11/20/2006
PATENT APPLICATION: **US/10/544,180A** TIME: 13:47:31

Input Set : **N:\efs\10544180a_efs\14-03_US_ST25.txt**
Output Set: **N:\CRF4\11202006\J544180A.raw**

<110> APPLICANT: Mohnen, Debra
Hahn, Michael G.
Kolli, Venkata S.K.
Doong, Ron L.
Sterling, Jason D.
<120> TITLE OF INVENTION: Galacturonosyltransferases, nucleic acids encoding same, and uses
therefor
<130> FILE REFERENCE: 14-03
<140> CURRENT APPLICATION NUMBER: US 10/544,180A
<141> CURRENT FILING DATE: 2006-03-08
<150> PRIOR APPLICATION NUMBER: US 60/445,539
<151> PRIOR FILING DATE: 2003-02-06
<150> PRIOR APPLICATION NUMBER: PCT/US04/03545
<151> PRIOR FILING DATE: 2004-02-05
<160> NUMBER OF SEQ ID NOS: 58
<170> SOFTWARE: PatentIn version 3.3
<210> SEQ ID NO: 1
<211> LENGTH: 2022
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis thaliana
<400> SEQUENCE: 1

Cpg-6)

atggcgctaa agcgagggct atctggagtt aaccggatta gaggaagtgg tggtggatct 60
cgatctgtgc ttgtgcttct catatttttc tgtgtttttg cacctctttg cttctttgtt 120
ggccgaggag tgtatatcga ttcctcaaat gattattcaa ttgtttctgt gaagcagaat 180
cttgactgga gagaacgttt agcaatgcaa tctgttagat ctcttttctc gaaagagata 240
ctagatgtta tagcaaccag cacagctgat ttgggtcctc ttagccttga ttcttttaag 300
aaaaacaatt tgtctgcatc atggcgggga accggagtag acccctcctt tagacattct 360
gagaatccag caactcctga tgtcaaatct aataacctga atgaaaaacg tgacagcatt 420
tcaaaagata gtatccatca gaaagttgag acacctacaa agattcacag aaggcaacta 480
agagagaaaa ggcgtgagat gcgggcaaat gagttagttc agcacaatga tgacacgatt 540
ttgaaactcg aaaatgctgc cattgaacgc tctaagtctg ttgattctgc agtccttggt 600
aaatacagta tttggagaag agaaaatgag aatgacaact ctgattcaaa tatacgcttg 660
atgcgggatc aagtaataat ggctagagtc tatagtggga ttgcaaaatt gaaaaacaag 720
aacgatttgt tacaagaact ccaggcccga cttaaggaca gccaacgggt tttgggggaa 780
gcaacatctg atgctgatct tcctcggagt gcgcatgaga aactcagagc catgggtcaa 840
gtcttggcta aagctaagat gcagttatat gactgcaagc tggttactgg aaagctgaga 900
gcaatgcttc agactgccga cgaacaagtg aggagcttaa agaagcagag tacttttctg 960
gctcagttag cagcaaaaac cattccaaat cctatccatt gcctatcaat gcgcttgact 1020
atcgattact atcttctgtc tccggagaaa agaaaattcc ctcggagtga aaacctagaa 1080
aaccctaatc tttatcatta tgccctcttt tccgacaatg tattagctgc atcagtagtt 1140
gttaactcaa ccatcatgaa tgccaaggat ccttctaagc atgtttttca ccttgtcacg 1200
gataaactca atttcggagc aatgaacatg tggttcctcc taaacccacc cggaaaggca 1260
accatacatg tggaaaacgt cgatgagttt aagtggctca attcatctta ctgtcctgtc 1320

RAW SEQUENCE LISTING DATE: 11/20/2006
PATENT APPLICATION: **US/10/544,180A** TIME: 13:47:31

Input Set : **N:\efs\10544180a_efs\14-03_US_ST25.txt**
Output Set: **N:\CRF4\11202006\J544180A.raw**

77 cttcgtcagc ttgaatctgc agcaatgaga gagtactatt ttaaagcaga ccatccaact 1380
79 tcaggctctt cgaatctaaa atacagaaac ccaaagtatc tatccatgtt gaatcacttg 1440
81 agattctacc tccctgaggt ttatcccaag ctgaacaaaa tcctcttcct ggacgatgac 1500
83 atcattgttc agaaagactt gactccactc tgggaagtta acctgaacgg caaagtcaac 1560
85 ggtgcagtcg aaacctgtgg ggaaagtttc cacagattcg acaagtatct caacttttcg 1620
87 aatcctcaca ttgcgaggaa cttcaatcca aatgcttgtg gatgggctta tggaatgaac 1680
89 atgttcgacc taaaggaatg gaagaagaga gacatcactg gtatatacca caagtggcaa 1740
91 aacatgaatg agaacaggac actatggaag ctagggacat tgccaccagg attaataaca 1800
93 ttctacggat taacacatcc cttaaacaag gcgtggcatg tgctgggact tggatataac 1860
95 ccgagtatcg acaagaagga cattgagaat gcagcagtgg ttcactataa cgggaacatg 1920
97 aaaccatggt tggagttggc aatgtccaaa tatcggccgt attggaccaa gtacatcaag 1980
99 tttgatcacc catatcttcg tcgttgcaac cttcatgaat aa 2022

102 <210> SEQ ID NO: 2
103 <211> LENGTH: 673
104 <212> TYPE: PRT
105 <213> ORGANISM: Arabidopsis thaliana
107 <400> SEQUENCE: 2

```
109 Met Ala Leu Lys Arg Gly Leu Ser Gly Val Asn Arg Ile Arg Gly Ser
110 1               5                  10                  15
113 Gly Gly Gly Ser Arg Ser Val Leu Val Leu Leu Ile Phe Phe Cys Val
114             20                  25                  30
117 Phe Ala Pro Leu Cys Phe Phe Val Gly Arg Gly Val Tyr Ile Asp Ser
118         35                  40                  45
121 Ser Asn Asp Tyr Ser Ile Val Ser Val Lys Gln Asn Leu Asp Trp Arg
122     50                  55                  60
125 Glu Arg Leu Ala Met Gln Ser Val Arg Ser Leu Phe Ser Lys Glu Ile
126 65                  70                  75                  80
129 Leu Asp Val Ile Ala Thr Ser Thr Ala Asp Leu Gly Pro Leu Ser Leu
130                 85                  90                  95
133 Asp Ser Phe Lys Lys Asn Asn Leu Ser Ala Ser Trp Arg Gly Thr Gly
134             100                 105                 110
137 Val Asp Pro Ser Phe Arg His Ser Glu Asn Pro Ala Thr Pro Asp Val
138         115                 120                 125
141 Lys Ser Asn Asn Leu Asn Glu Lys Arg Asp Ser Ile Ser Lys Asp Ser
142     130                 135                 140
145 Ile His Gln Lys Val Glu Thr Pro Thr Lys Ile His Arg Arg Gln Leu
146 145                 150                 155                 160
149 Arg Glu Lys Arg Arg Glu Met Arg Ala Asn Glu Leu Val Gln His Asn
150                 165                 170                 175
153 Asp Asp Thr Ile Leu Lys Leu Glu Asn Ala Ala Ile Glu Arg Ser Lys
154             180                 185                 190
157 Ser Val Asp Ser Ala Val Leu Gly Lys Tyr Ser Ile Trp Arg Arg Glu
158         195                 200                 205
161 Asn Glu Asn Asp Asn Ser Asp Ser Asn Ile Arg Leu Met Arg Asp Gln
162     210                 215                 220
165 Val Ile Met Ala Arg Val Tyr Ser Gly Ile Ala Lys Leu Lys Asn Lys
166 225                 230                 235                 240
169 Asn Asp Leu Leu Gln Glu Leu Gln Ala Arg Leu Lys Asp Ser Gln Arg
170                 245                 250                 255
```

RAW SEQUENCE LISTING DATE: 11/20/2006
PATENT APPLICATION: **US/10/544,180A** TIME: 13:47:31

Input Set : **N:\efs\10544180a_efs\14-03_US_ST25.txt**
Output Set: **N:\CRF4\11202006\J544180A.raw**

```
173 Val Leu Gly Glu Ala Thr Ser Asp Ala Asp Leu Pro Arg Ser Ala His
174             260                 265                 270
177 Glu Lys Leu Arg Ala Met Gly Gln Val Leu Ala Lys Ala Lys Met Gln
178         275                 280                 285
181 Leu Tyr Asp Cys Lys Leu Val Thr Gly Lys Leu Arg Ala Met Leu Gln
182     290                 295                 300
185 Thr Ala Asp Glu Gln Val Arg Ser Leu Lys Lys Gln Ser Thr Phe Leu
186 305                 310                 315                 320
189 Ala Gln Leu Ala Ala Lys Thr Ile Pro Asn Pro Ile His Cys Leu Ser
190                 325                 330                 335
193 Met Arg Leu Thr Ile Asp Tyr Tyr Leu Leu Ser Pro Glu Lys Arg Lys
194             340                 345                 350
197 Phe Pro Arg Ser Glu Asn Leu Glu Asn Pro Asn Leu Tyr His Tyr Ala
198         355                 360                 365
201 Leu Phe Ser Asp Asn Val Leu Ala Ala Ser Val Val Val Asn Ser Thr
202     370                 375                 380
205 Ile Met Asn Ala Lys Asp Pro Ser Lys His Val Phe His Leu Val Thr
206 385                 390                 395                 400
209 Asp Lys Leu Asn Phe Gly Ala Met Asn Met Trp Phe Leu Leu Asn Pro
210                 405                 410                 415
213 Pro Gly Lys Ala Thr Ile His Val Glu Asn Val Asp Glu Phe Lys Trp
214             420                 425                 430
217 Leu Asn Ser Ser Tyr Cys Pro Val Leu Arg Gln Leu Glu Ser Ala Ala
218         435                 440                 445
221 Met Arg Glu Tyr Tyr Phe Lys Ala Asp His Pro Thr Ser Gly Ser Ser
222     450                 455                 460
225 Asn Leu Lys Tyr Arg Asn Pro Lys Tyr Leu Ser Met Leu Asn His Leu
226 465                 470                 475                 480
229 Arg Phe Tyr Leu Pro Glu Val Tyr Pro Lys Leu Asn Lys Ile Leu Phe
230                 485                 490                 495
233 Leu Asp Asp Asp Ile Ile Val Gln Lys Asp Leu Thr Pro Leu Trp Glu
234             500                 505                 510
237 Val Asn Leu Asn Gly Lys Val Asn Gly Ala Val Glu Thr Cys Gly Glu
238         515                 520                 525
241 Ser Phe His Arg Phe Asp Lys Tyr Leu Asn Phe Ser Asn Pro His Ile
242     530                 535                 540
245 Ala Arg Asn Phe Asn Pro Asn Ala Cys Gly Trp Ala Tyr Gly Met Asn
246 545                 550                 555                 560
249 Met Phe Asp Leu Lys Glu Trp Lys Lys Arg Asp Ile Thr Gly Ile Tyr
250                 565                 570                 575
253 His Lys Trp Gln Asn Met Asn Glu Asn Arg Thr Leu Trp Lys Leu Gly
254             580                 585                 590
257 Thr Leu Pro Pro Gly Leu Ile Thr Phe Tyr Gly Leu Thr His Pro Leu
258         595                 600                 605
261 Asn Lys Ala Trp His Val Leu Gly Leu Gly Tyr Asn Pro Ser Ile Asp
262     610                 615                 620
265 Lys Lys Asp Ile Glu Asn Ala Ala Val Val His Tyr Asn Gly Asn Met
266 625                 630                 635                 640
269 Lys Pro Trp Leu Glu Leu Ala Met Ser Lys Tyr Arg Pro Tyr Trp Thr
```

RAW SEQUENCE LISTING DATE: 11/20/2006
PATENT APPLICATION: US/10/544,180A TIME: 13:47:31

Input Set : N:\efs\10544180a_efs\14-03_US_ST25.txt
Output Set: N:\CRF4\11202006\J544180A.raw

```
270                 645                 650                 655
273 Lys Tyr Ile Lys Phe Asp His Pro Tyr Leu Arg Arg Cys Asn Leu His
274             660                 665                 670
277 Glu
281 <210> SEQ ID NO: 3
282 <211> LENGTH: 1860
283 <212> TYPE: DNA
284 <213> ORGANISM: Arabidopsis thaliana
286 <400> SEQUENCE: 3
287 atgaaaggcg gaggcggtgg tggaggaggt ggtggcggag gaaaacgccg gtggaaagtt     60
289 ctggtgattg gagttttggt tcttgttatt ctttctatgc ttgttcctct tgctttctta    120
291 ctcggtcttc acaatggctt tcactctcct ggatttgtca ctgttcaacc ggcttcttca    180
293 tttgagagct ttaccagaat caatgctact aagcatacac agagagatgt atccgaacgg    240
295 gtcgatgagg ttcttcaaaa aatcaatcca gttcttccca agaaaagcga cataaacgtg    300
297 ggttccagag atgtgaatgc aacaagcggc actgattcta aaaaaagagg attaccagtg    360
299 tccccaactg ttgttgccaa tccaagccct gcaaataaaa caaaatcgga agcctcatat    420
301 acaggtgttc agaggaaaat agtaagtggt gatgaaactt ggagaacttg tgaagtgaaa    480
303 tatgggagct actgcctctg gagggaggaa aataaggaac caatgaaaga tgccaaggtg    540
305 aagcaaatga aggaccagct gtttgtggct agagcatact atcccagtat tgctaaaatg    600
307 ccttctcaaa gcaagttgac tcgggatatg aaacagaata tccaagagtt tgagcgtatt    660
309 cttagtgaaa gttctcaaga tgctgacctt ccaccacagg ttgataaaaa gttgcagaag    720
311 atggaagctg taattgcaaa ggcaaagtct tttccagtcg actgtaacaa tgttgacaag    780
313 aaattgagac agatccttga tttgactgag gatgaagcta gtttccacat gaaacagagt    840
315 gtgttcctct accagcttgc agtacagaca atgcctaaga gtcttcattg cttgtcaatg    900
317 cgactaactg tggaacattt caagtcagat tcacttgagg atcccattag tgagaaattt    960
319 tcagatccct cattacttca ctttgttatc atctccgata atatactagc atcgtccgtt   1020
321 gtgatcaact caacggttgt acatgcaagg gacagtaaaa actttgtttt ccatgtactg   1080
323 acagacgagc agaattactt tgcaatgaaa caatggttta ttaggaatcc ttgcaaacaa   1140
325 tcaactgttc aagtattgaa cattgaaaaa ctcgagctgg acgattctga tatgaaactg   1200
327 tctttgtctg cggagttccg tgtttccttc cccagtggtg acctttggc gtctcaacag   1260
329 aatagaacac actacttatc ccttttctct caatctcact atcttcttcc caaattattt   1320
331 gacaaattgg agaaggttgt gattctggat gatgacgttg tagtccagcg agacttatct   1380
333 ccccctttggg accttgatat ggaagggaaa gtgaatggcg ctgttaagtc gtgcactgtg   1440
335 agattgggtc agctaaggag tctcaagaga ggaaattttg ataccaatgc ttgtctctgg   1500
337 atgtctggtt tgaatgtcgt tgatcttgct agatggaggg cattgggtgt ttcagaaacc   1560
339 tatcaaaaat attataaaga gatgagtagt ggagatgagt cgagcgaagc aattgcattg   1620
341 caggcaagct tgctcacatt tcaagaccaa gtatatgctc ttgacgacaa atgggctcta   1680
343 tcagggcttg gttatgacta ctacatcaat gcacaagcca taaaaaacgc agccatattg   1740
345 cactataacg ggaacatgaa gccgtggctt gagctgggaa tcccaaatta caaaaactat   1800
347 tggagaaggc atctgagtcg ggaagatcgg ttcttgagtg actgtaacgt gaatccttga   1860
350 <210> SEQ ID NO: 4
351 <211> LENGTH: 619
352 <212> TYPE: PRT
353 <213> ORGANISM: Arabidopsis thaliana
355 <400> SEQUENCE: 4
357 Met Lys Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Lys Arg
358 1               5                  10                  15
361 Arg Trp Lys Val Leu Val Ile Gly Val Leu Val Leu Val Ile Leu Ser
362             20                  25                  30
```

RAW SEQUENCE LISTING DATE: 11/20/2006
PATENT APPLICATION: **US/10/544,180A** TIME: 13:47:31

Input Set : **N:\efs\10544180a_efs\14-03_US_ST25.txt**
Output Set: **N:\CRF4\11202006\J544180A.raw**

```
365 Met Leu Val Pro Leu Ala Phe Leu Leu Gly Leu His Asn Gly Phe His
366         35                  40                  45
369 Ser Pro Gly Phe Val Thr Val Gln Pro Ala Ser Ser Phe Glu Ser Phe
370     50                  55                  60
373 Thr Arg Ile Asn Ala Thr Lys His Thr Gln Arg Asp Val Ser Glu Arg
374 65                  70                  75                  80
377 Val Asp Glu Val Leu Gln Lys Ile Asn Pro Val Leu Pro Lys Lys Ser
378                 85                  90                  95
381 Asp Ile Asn Val Gly Ser Arg Asp Val Asn Ala Thr Ser Gly Thr Asp
382             100                 105                 110
385 Ser Lys Lys Arg Gly Leu Pro Val Ser Pro Thr Val Val Ala Asn Pro
386         115                 120                 125
389 Ser Pro Ala Asn Lys Thr Lys Ser Glu Ala Ser Tyr Thr Gly Val Gln
390     130                 135                 140
393 Arg Lys Ile Val Ser Gly Asp Glu Thr Trp Arg Thr Cys Glu Val Lys
394 145                 150                 155                 160
397 Tyr Gly Ser Tyr Cys Leu Trp Arg Glu Glu Asn Lys Glu Pro Met Lys
398                 165                 170                 175
401 Asp Ala Lys Val Lys Gln Met Lys Asp Gln Leu Phe Val Ala Arg Ala
402             180                 185                 190
405 Tyr Tyr Pro Ser Ile Ala Lys Met Pro Ser Gln Ser Lys Leu Thr Arg
406         195                 200                 205
409 Asp Met Lys Gln Asn Ile Gln Glu Phe Glu Arg Ile Leu Ser Glu Ser
410     210                 215                 220
413 Ser Gln Asp Ala Asp Leu Pro Pro Gln Val Asp Lys Lys Leu Gln Lys
414 225                 230                 235                 240
417 Met Glu Ala Val Ile Ala Lys Ala Lys Ser Phe Pro Val Asp Cys Asn
418                 245                 250                 255
421 Asn Val Asp Lys Lys Leu Arg Gln Ile Leu Asp Leu Thr Glu Asp Glu
422             260                 265                 270
425 Ala Ser Phe His Met Lys Gln Ser Val Phe Leu Tyr Gln Leu Ala Val
426         275                 280                 285
429 Gln Thr Met Pro Lys Ser Leu His Cys Leu Ser Met Arg Leu Thr Val
430     290                 295                 300
433 Glu His Phe Lys Ser Asp Ser Leu Glu Asp Pro Ile Ser Glu Lys Phe
434 305                 310                 315                 320
437 Ser Asp Pro Ser Leu Leu His Phe Val Ile Ile Ser Asp Asn Ile Leu
438                 325                 330                 335
441 Ala Ser Ser Val Val Ile Asn Ser Thr Val Val His Ala Arg Asp Ser
442             340                 345                 350
445 Lys Asn Phe Val Phe His Val Leu Thr Asp Glu Gln Asn Tyr Phe Ala
446         355                 360                 365
449 Met Lys Gln Trp Phe Ile Arg Asn Pro Cys Lys Gln Ser Thr Val Gln
450     370                 375                 380
453 Val Leu Asn Ile Glu Lys Leu Glu Leu Asp Asp Ser Asp Met Lys Leu
454 385                 390                 395                 400
457 Ser Leu Ser Ala Glu Phe Arg Val Ser Phe Pro Ser Gly Asp Leu Leu
458                 405                 410                 415
461 Ala Ser Gln Gln Asn Arg Thr His Tyr Leu Ser Leu Phe Ser Gln Ser
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 11/20/2006
PATENT APPLICATION: US/10/544,180A TIME: 13:47:32

Input Set : N:\efs\10544180a_efs\14-03_US_ST25.txt
Output Set: N:\CRF4\11202006\J544180A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:51; Xaa Pos. 2,3,4,5,8,9,10
Seq#:52; Xaa Pos. 2,3,4,5,6,8,9,10,11,12,13,15,17,18,19,21,23,24,25,27,28
Seq#:52; Xaa Pos. 29,30,31,32,33
Seq#:53; Xaa Pos. 2,3,4,5,6,7,8,9,10,11,12,13,14
Seq#:54; Xaa Pos. 2,3,4,5,6,9,10,12,15,16,18,19,20,21,22,23,24,25,26,27,28
Seq#:54; Xaa Pos. 30,32,33
Seq#:55; Xaa Pos. 2,4,5,7,9,10,13,15,16,17,18,19,20,21,22,23,25,26,28,29
Seq#:56; Xaa Pos. 3,4,5,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,26
Seq#:56; Xaa Pos. 27,29,30
Seq#:57; Xaa Pos. 2,4,5,6,7,8,10,11,12
Seq#:58; Xaa Pos. 2,3,4,5,6,7,9,10,11,12,13,14,15,16,17,18,20,21,22,26,27
Seq#:58; Xaa Pos. 28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,46,47
Seq#:58; Xaa Pos. 49,50

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:51,52,53,54,55,56,57,58

VERIFICATION SUMMARY DATE: 11/20/2006
PATENT APPLICATION: **US/10/544,180A** TIME: 13:47:32

Input Set : **N:\efs\10544180a_efs\14-03_US_ST25.txt**
Output Set: **N:\CRF4\11202006\J544180A.raw**

L:4754 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0
L:4774 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0
L:4778 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:16
L:4782 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:32
L:4802 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53 after pos.:0
L:4822 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0
L:4826 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:16
L:4830 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:32
L:4850 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0
L:4854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:16
L:4893 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0
L:4897 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:16
L:4921 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:0
L:4941 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:0
L:4945 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:16
L:4949 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:32
L:4953 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:48